



# Estimating rainfall distributions at high temporal resolutions using a multifractal model

A. Pathirana, S. Herath, T. Yamada

## ► To cite this version:

A. Pathirana, S. Herath, T. Yamada. Estimating rainfall distributions at high temporal resolutions using a multifractal model. *Hydrology and Earth System Sciences Discussions*, 2003, 7 (5), pp.668-679. hal-00304913

**HAL Id: hal-00304913**

**<https://hal.science/hal-00304913>**

Submitted on 1 Jan 2003

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Estimating rainfall distributions at high temporal resolutions using a multifractal model

Assela Pathirana<sup>1</sup>, Srikantha Herath<sup>1</sup> and Tadashi Yamada<sup>2</sup>

<sup>1</sup>Environment and Sustainable Development Program, United Nations University, 53-70 Jingumae 5-chome, Shibuya-ku, Tokyo 150-8925, Japan

<sup>2</sup>Faculty of Science and Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan

Email for corresponding author: asselapathirana@yahoo.co.jp

## Abstract

Rainfall data from 18 stations in the vicinity of Tokyo city, measured to a precision of 1 mm, were analysed for multifractal properties. A multifractal model based on the scaling properties of temporal distribution of rainfall intensities was formulated to investigate the intensity distribution relationships in the available scaling regime. Although conventional analysis did not provide encouraging results with these measurements, an alternative approach that could be applied to rainfall data of widely variable quality and duration was used to establish a scaling relationship between daily and hourly rainfall intensities. Using a discrete cascade algorithm based on the log-Lévy generator, synthetic hourly rainfall series were generated from the multifractal statistics of daily-accumulated rainfall. Several properties of rainfall time series that are relevant to the use of rainfall data in surface hydrological studies were used to determine, statistically, the degree of agreement between the synthetic hourly series and observed hourly rainfall.

**Keywords:** rainfall modelling, cascades, multifractal, downscaling

## Introduction

The relationship between rainfall intensities at different scales is not trivial owing to the high variability that rainfall exhibits in time and space, particularly at the scales that are of interest to the hydrologist. Rainfall intensities increase sharply in magnitude as the temporal accumulation length (time scale) is decreased. Thus, an understanding of the scaling relationship of rainfall intensity is an important requirement for assessing the effects of rainfall at smaller scales than those at which it is commonly observed. One important example is the utilisation of historical rainfall records to make statistical evaluation of catchment response: Long historical rainfall records are important resources for understanding the ‘behaviour’ of catchments, without excessive interference from possible annual or inter-annual anomalies. However, most historical data sets spanning more than a several decades are of daily observations — a fact that is especially true in the developing world. Development of techniques to estimate rainfall properties at higher resolutions such as hourly, based on daily records,

therefore is important in utilising these data resources effectively in response studies.

Since the introduction of the theory of fractals by Mandelbrot (1967), numerous attempts have been made to apply fractals to modelling various physical phenomena; Lovejoy and Mandelbrot (1985) applied them to studies of rainfall). However, Gupta and Waymire (1990), Lavallée *et al.* (1991), Lovejoy and Schertzer (1990) showed that many geophysical fields like cloud formations, temporal and spatial distributions of rainfall, river network formations, river flows and topography are modelled better by multiple scaling schemes.

Unlike single fractals, the multifractal models involve infinite numbers of parameters in general in their original forms and are of limited practical use. Several simplified forms, more useful in actual data analysis, have been proposed. In the ‘Universal Multifractal Theory’ proposed by Schertzer and Lovejoy (1987), the multifractal distributions are realised by a finite number of parameters based on the simplifying assumption that the multifractal

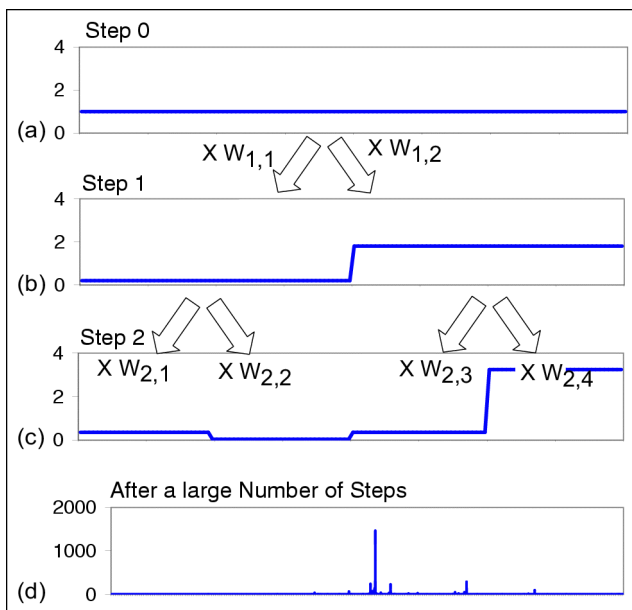


Fig. 1. A multiplicative cascade in one-dimension. A field of a fixed intensity is divided into two parts and each is multiplied by a 'weight'  $W_{i,j}$  which is derived from a log-Lévy random variable.

measure is a result of a special type of multiplicative cascade process. Figure 1 shows a 'discrete version' of a multiplicative cascade in one-dimension. Cascades had been widely used to model the energy transfer in turbulence (e.g. Novikov and Stewart, 1964; Mandelbrot, 1974; Frisch *et al.*, 1978). Schertzer and Lovejoy (1987) proposed a physical basis for the use of multiplicative cascade processes

in modelling atmospheric turbulence. A multiscaling approach has become the basis for statistical models that are widely used to model geophysical phenomena like rainfall and cloud formations, which are closely related to atmospheric turbulence (Lovejoy and Schertzer, 1990; Gupta and Waymire, 1990; Svensson *et al.*, 1996; de Lima and Grasman, 1999; Deidda *et al.*, 1999).

Many of the multifractal analyses and especially modelling efforts on rainfall process had been carried out with relatively high quality rainfall data, which are available under very special circumstances (see Olsson, 1996; Svensson *et al.*, 1996; Olsson, 1998; de Lima and Grasman, 1999, for some examples). While being extremely useful in understanding the multifractal nature of rainfall and in validating multifractal models proposed to describe the scaling of rainfall process, these approaches, when applied to rainfall data of typical accuracy and durations, often produce discouraging results, particularly in the case of modelling high resolution distributions from observations at a lower resolution.

The calculation scheme described here makes it possible to use a multifractal model to predict hourly rainfall properties successfully from daily rainfall observations. First, the details of the data analysis are presented. Hourly rainfall records over two decades from 18 raingauges around Tokyo (Fig. 2 and Table 1) were selected for the study. These were measured by tipping bucket raingauges of 1mm precision. Daily rainfall series, obtained by accumulating those hourly records, were used as source data to fit the

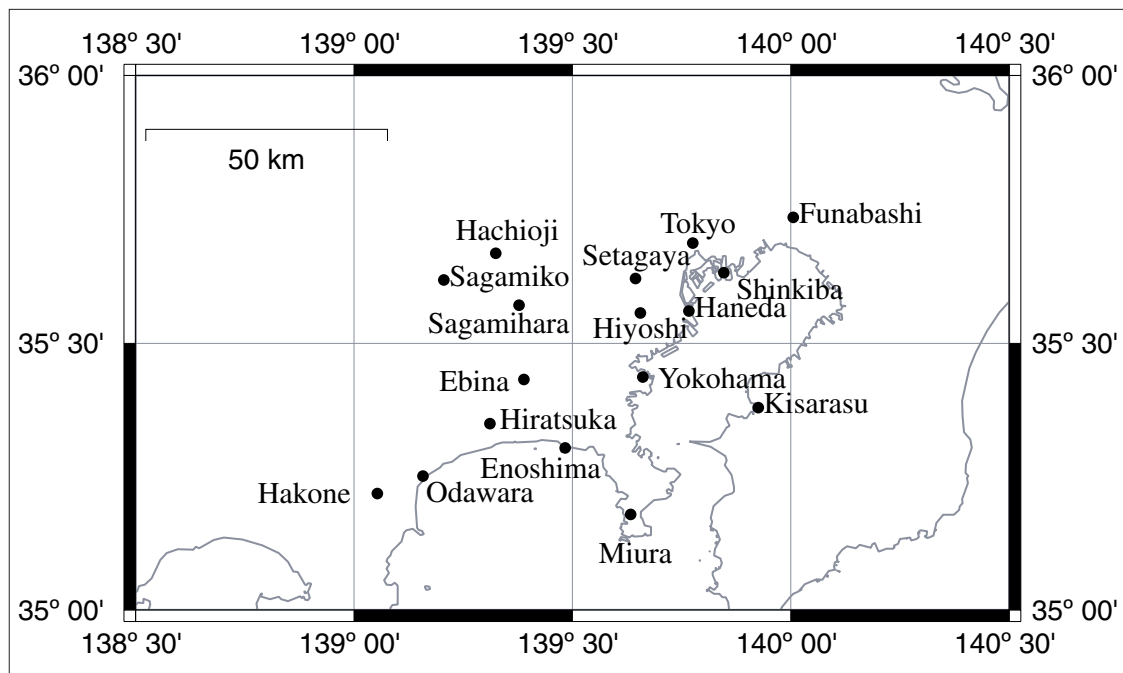


Fig. 2. Locations of the selected rain gauges.

Table 1. Some statistical and multifractal properties of the selected rainfall series.

	Elevation (m)	Average rainfall (mm/yr)	Series length (yrs)	$C_1$	$H$	$\alpha$	$\beta$
Ebina	18	1702	22	0.35	-0.06	1.34	1.05
Enoshima	60	1551	17	0.32	-0.04	1.35	1.00
Funabashi	24	1295	20	0.32	-0.05	1.37	0.84
Hachioji	121	1570	8	0.40	-0.07	1.09	1.11
Hakone	850	3421	22	0.31	-0.06	1.47	1.21
Haneda	3	1374	22	0.33	-0.05	1.37	0.99
Hiratsuka	20	1574	22	0.35	-0.06	1.45	1.01
Hiyoshi	57	1510	22	0.31	-0.05	1.63	1.07
Kisarasu	5	1453	22	0.35	-0.05	1.21	0.90
Miura	42	1567	17	0.35	-0.06	1.42	1.05
Odawara	28	1968	22	0.39	-0.05	1.09	1.12
Sagamihara	149	1677	22	0.38	-0.06	1.29	1.05
Sagamiko	188	1552	22	0.42	-0.08	1.23	1.11
Setagaya	41	1526	22	0.35	-0.06	1.33	0.88
Shinkiba	6	1311	22	0.34	-0.05	1.31	0.96
Tokyo	7	1450	22	0.36	-0.05	1.19	1.03
Yokohama	39	1625	22	0.26	-0.04	1.92	1.06
Average				0.345	-0.056	1.357	1.026
St.dev.				0.038	0.009	0.198	0.094

multifractal model. Then the next section explains the procedure of generating synthetic hourly rainfall series based on that model. Finally, several statistical properties that are relevant for the use of rainfall data for surface hydrology, were used to compare the synthetic rainfall series with actual hourly observations, the results are discussed and conclusions are arrived at.

## Data Analysis

### UNIVERSAL MULTIFRACTALS

A field  $\phi$  is said to be *conserved* if the overall mean of the field ( $\langle\phi\rangle$ ) is independent of the scale  $t$ . In the formulation of the Probability Distribution Multiple Scaling (PDMS) method (Lovejoy and Schertzer, 1990), where data are considered continuous rather than as a set, the properties of a conserved multiscaling field  $\phi_\lambda$  are represented by the following probabilistic relation (Tessier *et al.*, 1993):

$$P(\phi_\lambda \geq \lambda^\gamma) \approx \lambda^{-c(\gamma)} \quad (1)$$

where scale ratio  $\lambda$  is obtained by dividing the largest scale of interest  $T$  ( $T=1024$  h was selected for the present analysis and this selection is discussed later.) by the scale  $t$ .  $\phi$  is the field obtained by normalising the original field by the overall

mean of the field at the highest resolution (largest  $\lambda$ ) of interest.  $\gamma$  is a scaling exponent and  $c(\gamma)$  is known as the *codimension function*. (Although this relationship is sometimes written as an exact equation in recent literature, this relationship is an approximation valid only for “slowly varying quantities like logarithms” (Tessier *et al.*, 1993) and, thus, cannot be treated as an exact equality.) Assuming that the underlying cascade process is based on a Lévy-stable distribution, the codimension function can be shown to have the following functional form (Tessier *et al.*, 1993), which reduces the infinite number of parameters in the general multifractal case, to three:

$$c(\gamma - H) = \begin{cases} C_1 \left( \frac{\gamma}{C_1 \alpha} + \frac{1}{\alpha} \right)^\alpha & \alpha \neq 1 \\ C_1 \exp\left(\frac{\gamma}{C_1} - 1\right) & \alpha = 1 \end{cases} \quad (2)$$

(for  $0 \leq \alpha \leq 2$ )

where  $1/\alpha + 1/\alpha' = 1$ .  $H$  is the exponent of the power-law filter in the frequency domain required to obtain  $\phi$  the non-conserved (observed) field  $R_\lambda$  ( $R_\lambda \approx \phi_\lambda \lambda^{-H}$ ).  $C_1$  is the value of the codimension for the mean process. The multifractality parameter  $\alpha$ , also known as the Lévy index, indicates the probability distribution involved in the underlying multiplicative cascade process.  $\alpha = 0$  is for the monofractal case and values larger than zero indicate

multifractals. The case  $\alpha=2$  corresponds to multifractals with lognormal, or Gaussian generators.

Two important properties of the codimension curve (the plot of  $c(\gamma)$  against  $\gamma$ ) are: (1) it is a convex function for  $\alpha > 0$ ; (2) the tangent drawn to the curve at the mean process ( $c(\gamma) = C_1$ ) should be parallel to the bisectrix,  $c(\gamma) = \gamma$  (it passes through the origin in case of a conserved process,  $H = 0$ ).

## METHODOLOGY

The widely used method (based on descriptions by Lavallée *et al.*, 1991 and Olsson *et al.*, 1993) for calculation of the codimension function is as follows: the multifractal theory proposes rainfall process to be a scale-independent phenomenon. However, rainfall data are measured (or processed later as in the case of tipping bucket data) at some specific resolution (temporal in the present context), which will be referred to as 'measured-scale' hereafter. Firstly, the data will be normalised at the measured-scale, by dividing each intensity value by the average rainfall intensity, so that the average value of the rainfall series is unity. The rainfall series is then analysed at different temporal scales,  $t$  (sometimes known as the 'box-size'). For example, in the case of data of hourly measured-scale, it is possible to analyse at  $t=1\text{h}$ ,  $2\text{h}$ ,  $24\text{h}$ , etc., i.e. any multiple of measured-scale can be analysed. However, due to the power-law nature of the relationship of  $c(\gamma)$  to the other terms in Eqn. (1), it is logical and statistically robust to follow a power series like  $2\text{h}$ ,  $4\text{h}$ ,  $8\text{h}$ ,  $16\text{h}$ , etc., unless there is a special reason to do otherwise.

To compute rainfall values at a scale  $t$ , starting from the beginning of the rainfall series, each consecutive  $n$  value ( $n = t/t_m$ , where  $t_m$  is the measured-scale) are added together. This will make a new series of length  $L = L_m / n$ , where  $L_m$  is the length of series at measured-scale. Then at scale  $t$ , the exceedance probability,  $P(\phi_\lambda \geq \lambda^\gamma)$  is computed for a number of  $\gamma$  values.

The applicable range of  $\gamma$  values for a particular analysis is determined by two factors: (1) The upper bound of  $\gamma$  values depends on the availability of high values of the measure, beyond some value of  $\gamma$  the  $\lambda^\gamma$  value becomes so large that, there almost are no boxes which satisfy the condition  $R_\lambda \geq \lambda^\gamma$  (see Tessier *et al.*, 1993, for a theoretical discussion on this upper bound). (2) At small values of  $\gamma$  (i.e. small  $\lambda^\gamma$ ), the effect of the precision of measurement becomes prominent. This becomes a problem especially at small box sizes (large  $\lambda$ ). Low precision of measurement can induce a step-like behaviour in (the lower values of) the exceedance probabilities.

The exceedance probability is approximated by  $N_\lambda(\gamma) /$

$N_\lambda$  where  $N_\lambda(\gamma)$  is the number of boxes whose average intensity exceeds  $\lambda^\gamma$  (where  $\lambda = T/t$ ), out of the total number of boxes,  $N_\lambda = L/t$ .

To test the validity of Eqn. (1),  $P(\phi_\lambda \geq \lambda^\gamma)$  is plotted against  $\lambda$  in a log-log domain. Figure 3 shows a number of such curves resulting from the standard application of PDMS method to the Tokyo dataset. (For this analysis, hourly observations at the original hourly resolution were used.) The straight regression lines obtained for each value of  $\gamma$  indicate the existence of a scaling regime covering resolutions from one hour to about two days. The values for the codimension function were estimated from the slope

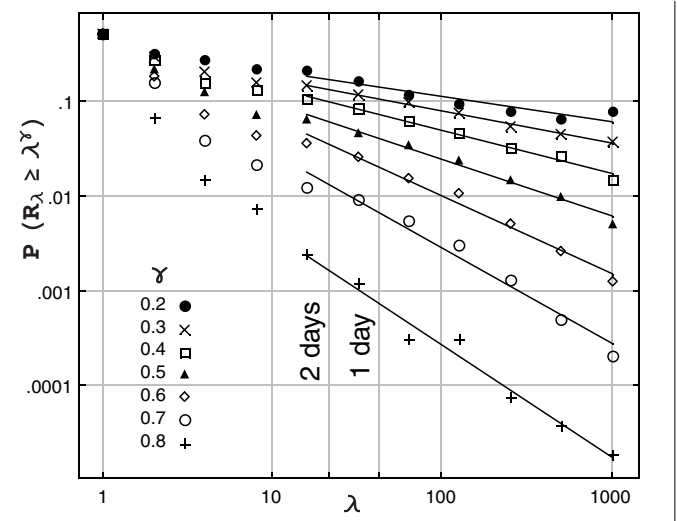


Fig. 3. The  $P - \lambda$  curves for the observations at Tokyo station. Only a limited number of  $\gamma$  values are shown in the figure for clarity. Ordinates corresponding to scales of 1 day and 2 days are marked on the graph.

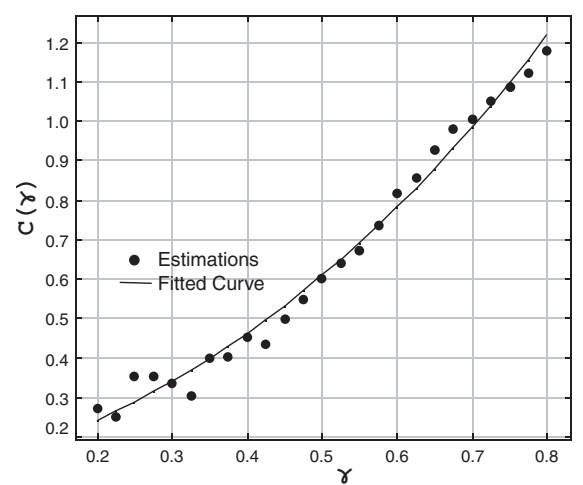


Fig. 4. The codimension function estimated for Tokyo. The scatter observed at low  $\gamma$  values may be attributed to the low measuring precision.

of the linear fit for each value of  $\gamma$ . A non-linear estimation method (methods proposed by Meyer and Roth, 1972 and Rosenbrock, 1960, work well for this problem) was used to estimate the parameters  $H$ ,  $C$  and  $\alpha$  of a fitted codimension function. The fitted codimension function for the example series is shown in Fig. 4, together with the values estimated from data.

## Modelling hourly rainfall from daily observations

To attain the intended objective of this paper, a multifractal model has to be calibrated from daily data alone. For a hypothetical case of perfect scaling behaviour, this will not be a problem since scales above one day will have the same scaling behaviour as those below. However, in general, rainfall does not always show such perfect scaling properties. It has been reported that rainfall shows ‘breaks’ in scaling around a scale of a few days (see Pathirana, 2001). One of the widely used methods to understand scaling is the power-law behaviour of the spectral density function of a field. Within the scaling regime, the power spectral density  $S(k)$  (defined as  $S(k) = \lim_{T \rightarrow 0} 1/T |Y(k, T)|^2$  where Fourier transform is denoted by  $Y(k, T)$ ) relates to the wave number,  $k$  as follows:

$$S(k) \sim k^{-\beta} \quad (3)$$

Figure 5 shows the power spectral density functions for the rainfall series selected. The spectra indicate that the scaling regime, present in smaller scales, does not continue beyond two days resolution. This break in scaling of temporal rainfall spectra has been widely reported. The precise level of scale breaking appears to be affected by the sampling properties of data sets such as measurement precision and resolution, in addition to the characteristics of the rainfall itself.

As in previously reported studies, the present analysis also indicates the existence of a continuous scaling regime from daily to hourly scale. However, the fundamental problem in using daily data to model hourly rainfall is that the early break in scaling causes the ‘available’ scales for fitting a multifractal model to become severely limited. In the case of the Tokyo data, only 24 h and 48 h scales are available in the required scaling regime. The PDMS method described earlier, as well as the other methods involving linear regression in log-log space (e.g. the double trace moment method proposed by Lavallée, 1991) fail to work due to the above limitation in the data.

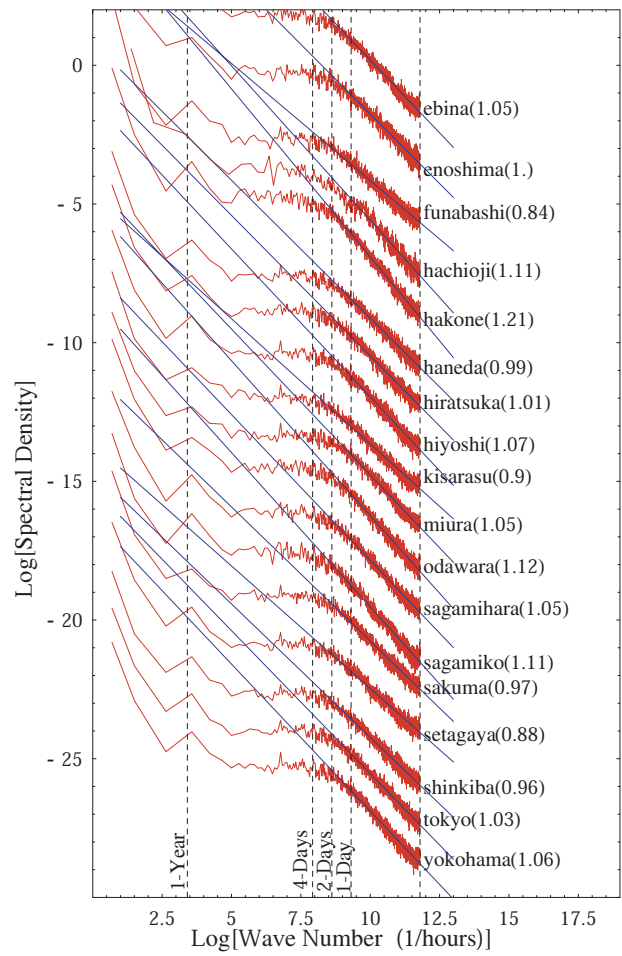


Fig. 5. Power spectral density plots for the rainfall series. In many cases the scaling breaks at a scale around 1-2 days. The spectral slopes are given within parentheses for each station. The data were smoothed by averaging over a variable interval proportional to the logarithm of wave number before plotting. The power spectral curves were shifted vertically.

### AN ALTERNATIVE APPROACH FOR ESTIMATING $c(\gamma)$

For a given  $\gamma$  value, the variation of  $\log(P(R \geq \lambda^\gamma))$  with the logarithm of non-dimensional scale,  $\lambda$ , can be represented by the following exact equation:

$$\log(P(R \geq \lambda^\gamma)) = -c(\gamma)\log(\lambda) + \log(b(\gamma)) \quad (4)$$

where  $b(\gamma)$  is a function independent of  $c(\gamma)$  and  $\lambda$ . Strictly,  $b(\gamma)$  can be a slowly varying function of the other variables, so that it ‘appears’ to be constant in the log-log space, for a given  $\gamma$ . In order to propose the alternative method,  $b$  is assumed to be a constant. The consequence of this assumption is illustrated in Fig. 6. All the  $\gamma = \text{constant}$  lines should converge at  $(\log[\lambda] = 0, P(R \geq \lambda^\gamma) = b)$ . Then, by rearranging,



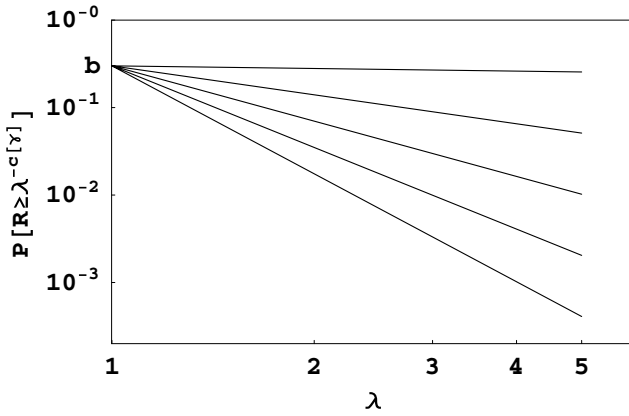


Fig. 6. The assumption of  $b(\gamma) \Rightarrow b$  requires the  $\gamma = \text{const.}$  lines to be converged on the y-axis ( $\lambda = 1$ ).

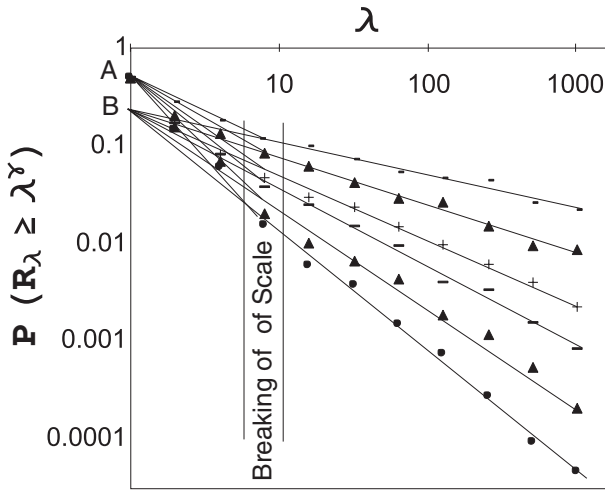


Fig. 7. When there are breaks in the scale, the value of  $b$  may not be equal to the exceedance probability of the mean of the field at the largest scale of interest ( $\lambda = 1$ ). While the point A indicates the exceedance probability of the mean of the field at  $\lambda = 1$ , the scaling regime in higher resolutions (larger  $\lambda$ ) has a  $b$  value corresponding to the ordinate at point B.

$$c(\gamma) = -\frac{\log(P(\phi_\lambda \geq \lambda^\gamma)/b)}{\log(\lambda)} \quad (5)$$

where  $b$  is a proportionality constant. For a hypothetical case of perfect scaling along the whole range of scales of interest (from largest  $\lambda$  to  $\lambda = 1$ ), one is compelled to derive a value for  $b$  by an obvious means: by definition, the value of  $\lambda$  for the largest scale of interest is 1. Hence,  $b = P(R_\lambda > 1)$  (from Eqn. 1). Since the original series is normalised (mean = 1) before analysis, this indicates the probability of exceedance of the mean of the field at the largest scale of interest.

However, most geophysical fields do not show a continuous scaling property all through the scale range of interest, but show one or more breaks in scale (Fig. 7.) So, it is generally not possible to estimate  $b$  by the above means

and a different empirical approach is used to find a suitable estimate for  $b$ .

#### ESTIMATING THE CONSTANT $B$

In the case of modelling with daily data only, the value of  $b$  is calculated as follows: since there are two scales (24 h and 48 h) available for analysis, given a value for  $b$ , two different estimations for  $c(\gamma)$  can be made as  $c_{b,24}(\gamma)$  and  $c_{b,48}(\gamma)$  using Eqn. 5. However, there are properties of  $c(\gamma)$  function that these estimates should satisfy, namely: (1) the estimations of  $c(\gamma)$  at different scales in the same scaling regime must be similar and (2) since rainfall is a conserved process, the bisectrix should be tangential to the codimension function (Tessier *et al.*, 1993). It was found that the above two properties are adequate to estimate a value for  $b$  using the following trial and error process: The values of  $c_{b,24}(\gamma)$  and  $c_{b,48}(\gamma)$  were estimated for various values of  $b$  until a value which satisfies both the above conditions approximately, is selected. Typical values of  $b$  were found to be around 0.30 to 0.40, for the rainfall series used. Figure 8 shows the estimation of  $b$  for Tokyo series.

#### SIMULATION OF MULTIFRACTAL FIELDS

There is a number of different multifractal simulation methods. Since the universal multifractal model is based on a Lévy-stable distribution, a cascade model with a log-Lévy process was used for the simulations. Two major variations of cascade schemes are used widely, namely, continuous cascade models and discrete cascade models. The former are sometimes claimed to be superior in maintaining the structure of rainfall, so that the data generated preserves indicators of the structure like spectra and autocorrelations more accurately. The main advantage of the discrete cascade models is that they are comparatively simpler and economical in computing. In the present research, a discrete cascade algorithm was used, based on the model described in Monin and Yaglom (1975), generalised to include a Lévy-stable distribution.

The model operation is simple and is based on the cascade model described in Fig. 1. The generation is started with a uniform field of unit intensity. The field is divided into two equal parts and each part is multiplied by a weight,  $W$ , that is derived from a log-Lévy random variable.

A Lévy random variable,  $L$ , with the following probability distribution was used in the model:

$$\begin{aligned} P(X > x) &\sim |x|^{-\alpha} \quad (0 < x) \\ P(X > x) &\sim \exp(-|x|^{\alpha'}) \quad (x < 0) \\ (0 < \alpha \leq 2) \text{ and } 1/\alpha + 1/\alpha' &= 1 \quad (6) \end{aligned}$$

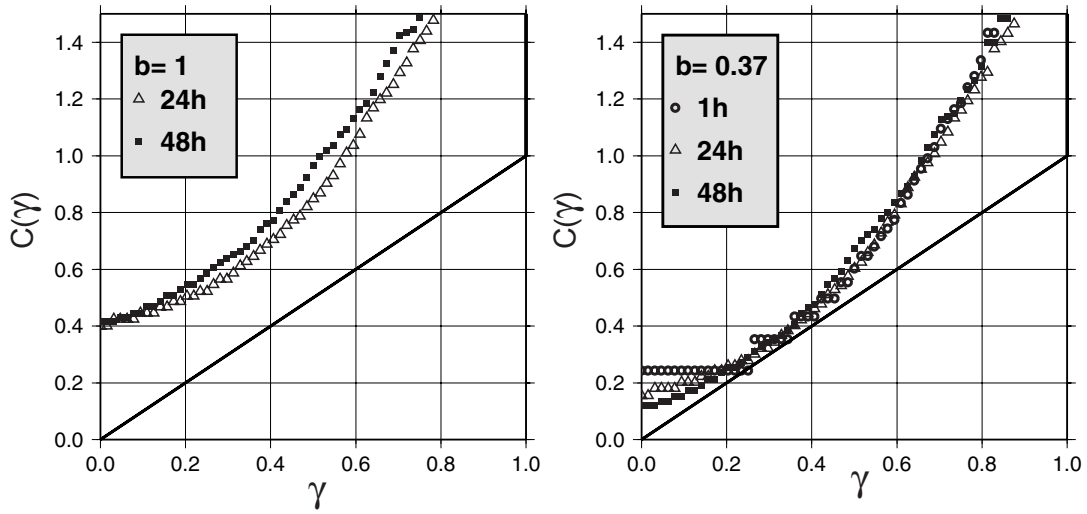


Fig. 8. Estimation of  $b$  and the computation of the codimension function are done together. Appropriate value of  $b$  is selected such that the codimension estimates at 24h and 48h are similar and the bisectrix is tangential to the estimate. Example: Funabashi. The trivial value,  $b = 1$  (left) does not work. A value of  $b = 0.37$  approximately fulfills the above conditions (right). Hourly estimate of the codimension function which is not involved in estimation, is also shown for the purpose of comparison.

While it is difficult to provide close-form expressions for the random variables of these distributions (with the exception of the  $\alpha = 2$  case where the distribution is log-normal), Lèvy-stable random variables can be generated numerically. Wilson *et al.* (1991) and Grigoriu (1995), among others have described numerical algorithms to generate Lèvy-stable random variables.

A value  $s$  drawn from a unit Lèvy random variable, should be modified as

$$L = s \left( \frac{C_1}{|1-\alpha|} \right)^{(1/\alpha)} \pm \frac{C_1}{|1-\alpha|} \text{ where } \pm \text{ signs for } \alpha \leq 1 \quad (7)$$

to obtain a Lèvy variable with proper magnitude and centering. The cascade weight is given by

$$W = \exp[L]. \quad (8)$$

This procedure makes sure that the expected value of the cascade weights,  $E(W) = 1$ .

One other variable that is important in analysis of data as well as generating rainfall series synthetically, is the selection of the largest scale of interest  $T$ . Trials with different values of  $T$  for data analysis using both PDMS (traditional method) and the new method proposed in this paper showed that the analysis depends critically on the selected  $T$  values.  $T=1024$  h worked well for analysing present data. Selection of values of  $T$  far from this (e.g.  $T$  smaller than 512 h or larger than 2048 h) caused the generated rainfall fields to have different intensity distributions than the observations, especially for extreme (rare) events, even after the proper

magnification given in Eqn. (9) was applied. In the simulation phase, it is important to use an appropriate number of cascade steps, depending on the selection of  $T$ . For example, to generate hourly data when  $T=1024$ , 10 cascade steps ( $\log_b(1024)$  where  $b$  is the number of branches, each value is divided at cascading, which is two for the present case) should be performed.

#### SYNTHETIC RAINFALLS FROM MULTIFRACTAL FIELDS

Once a multifractal field representing prescribed values of  $C_1$  and  $\alpha$  is obtained, two main steps have to be performed to convert it into a synthetic rainfall field. The generation process was started with a field of unit intensity and thus, the generated field's average value,  $\overline{M}$  should satisfy  $E(\overline{M}) = 1.0$ . To obtain a 'rain' series  $R$  of proper magnitude, the multifractal field is transformed as

$$R_i = M_i \overline{R} / \overline{M} \quad (9)$$

Where  $R_i$  is the rainfall value corresponding to the multifractal value  $M_i$ .  $\overline{M}$  is the average of the multifractal field generated (whose expected value is 1) and  $\overline{R}$  is the mean hourly rainfall intensity computed from daily rainfall observations.

The multifractal simulation model does not produce any zero values. The no-rain periods are created by implementing a lower cut-off limit to the intensities. However, the appropriate value of the lower cutoff can be decided only by comparing synthetic data with actual observations. In a situation where only daily data are



available, the value is decided by comparing the zero value fractions at the daily scale. This was done in the present study. Since the introduction of zero values reduces the average of the series by a small fraction (though in reality this reduction is negligibly small due to the inherent nature of multiplicative fields — the majority of low intensity values have negligible mass concentrations), magnitude adjustment was performed after introducing the zero values.

## Results

Firstly, the consistency and accuracy of the model was tested by investigating the spectra and codimension functions. The model-generated synthetic data showed similar multifractal scaling properties to the observed series (Fig. 9 shows the spectrum and codimension functions for synthetically generated rainfall for Funabashi station.) The spectra of synthetic data also show a break, though the spectra at smaller frequencies are less flat than those of the observations.

The comparison of Quantile-Quantile plots of data to evaluate the hourly intensity prediction performance of the model is shown in Fig. 10 for six of the 17 rainfall series analysed. To obtain a statistically meaningful comparison, the whole set of non-zero intensity values in the observed series, divided into an equal number of quantiles (400), was compared with the estimated intensities for the same quantiles.

While the overall agreement between observed and generated distributions is satisfactory, there is some deviation towards the high-intensity end. The estimation error at the last quantile varies from a virtually negligible level (e.g. Funabashi) to about 30% in the case of Hakone. With a few exceptions, most of the deviations at the high intensities observed in this study were towards overestimating the intensities by the generated distributions.

Comparison of the fractal dimension,  $d$ , (Feder, 1989) provides important clues to the structure of occurrence/non-occurrence of rainfall at a range of scales. A comparison of the wet-fractions at different scales, using the fractal scaling theory, is presented in Fig. 11. The agreement at different scales indicates that the amounts as well as the structure of the wet fractions are in close agreement.

The autocorrelation functions for some data sets are shown in Fig. 12. The observed and generated values agree closely at small lags (up to about 6 or 7 h). However, there is a residual amount of autocorrelation left in the synthetic data at larger lags. Though it was not possible to find a definite explanation for this spurious autocorrelation at long lags, the following is a likely reason: the observed data show a break in scaling (indicated both by a flattening down of the

spectra and breaks in  $P - \lambda$  curves) at scales above two days. However, since such imperfection in scaling was not considered explicitly in the cascade scheme, it produces fields with perfect scaling at all scales and as a result has a larger correlation at coarser scales than observed data.

The reason for this extended autocorrelation structure of synthetic data may be the implicit assumption that the rainfall process is perfectly scaled at the simulation stage, while actual data showed a break of scaling around two days.

For hydrological response studies such as flood simulations, the properties of whole rainfall events are also

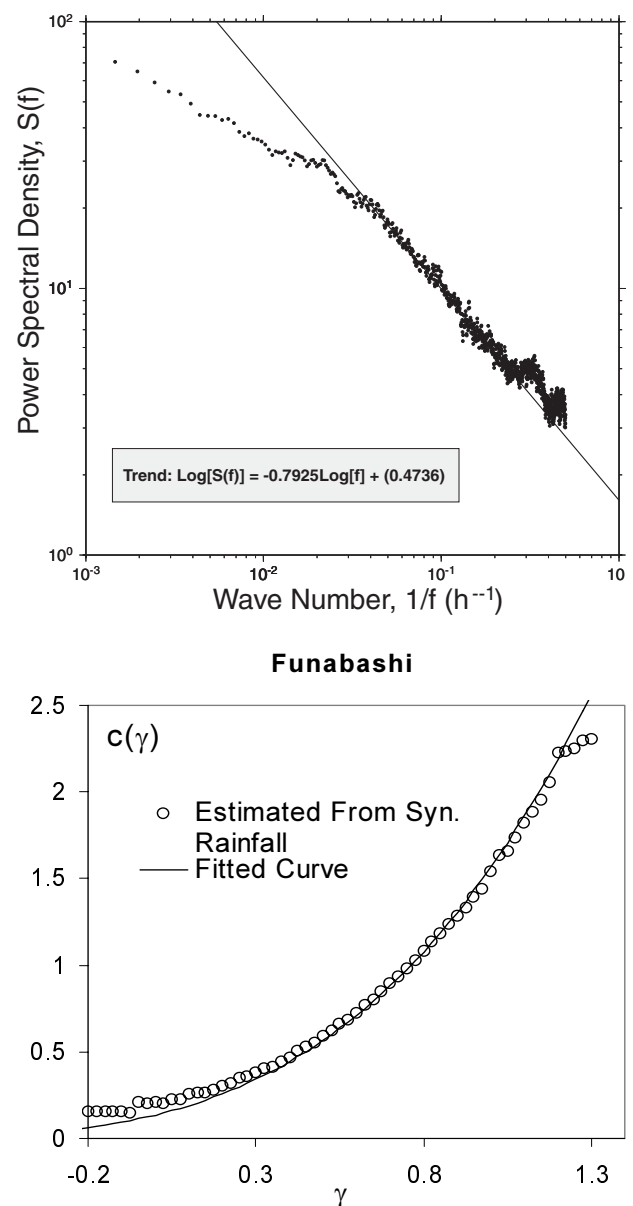


Fig. 9. Spectrum (top) and codimension function (bottom) of generated series for Funabashi station.

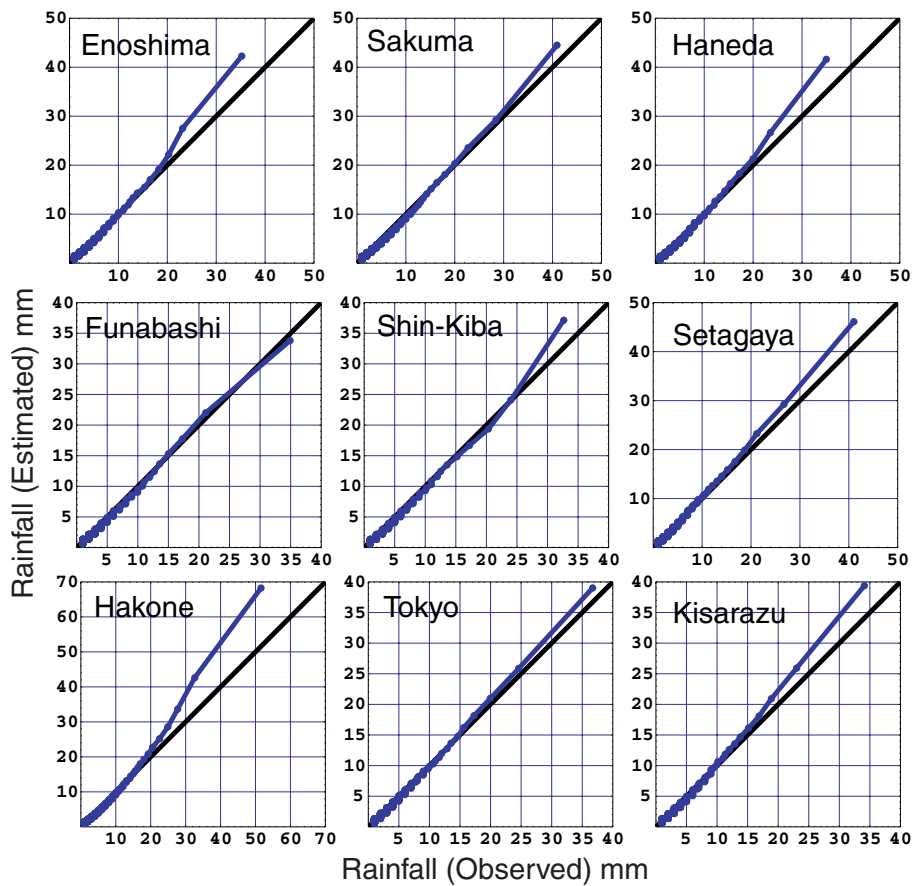


Fig. 10. Quantile-Quantile plots for rainfall distributions. Horizontal axis shows the values for the observed 1hr rainfall and the vertical axis shows the corresponding estimates using the daily rainfall distributions.

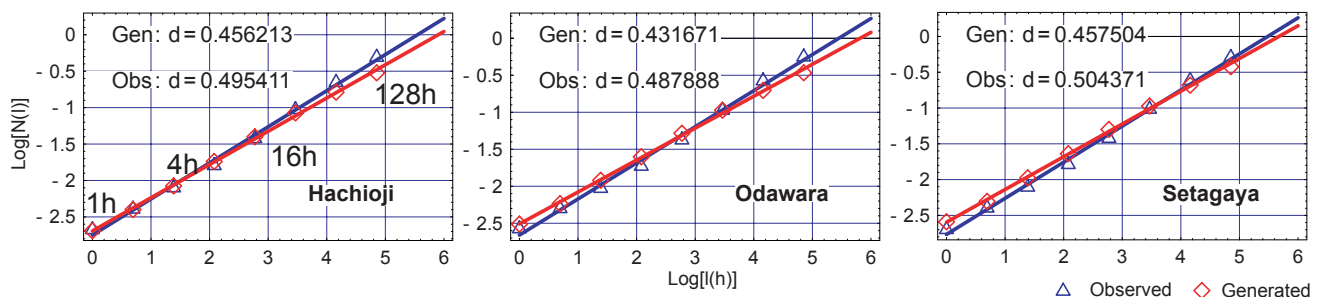


Fig. 11. Amount and distribution of non-zero values of three selected series.  $l$ : The 'box-size' or the resolution, in hours.  $N(l)$ : The number of boxes with non-zero rainfall. Agreement of fractal dimensions ( $d$ ) show that the structures (distribution) of the rainy periods are matching.

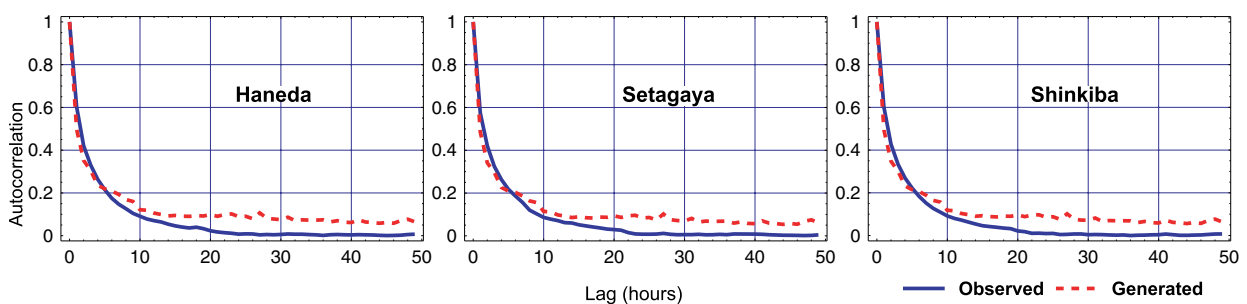


Fig. 12. Comparison of autocorrelations.

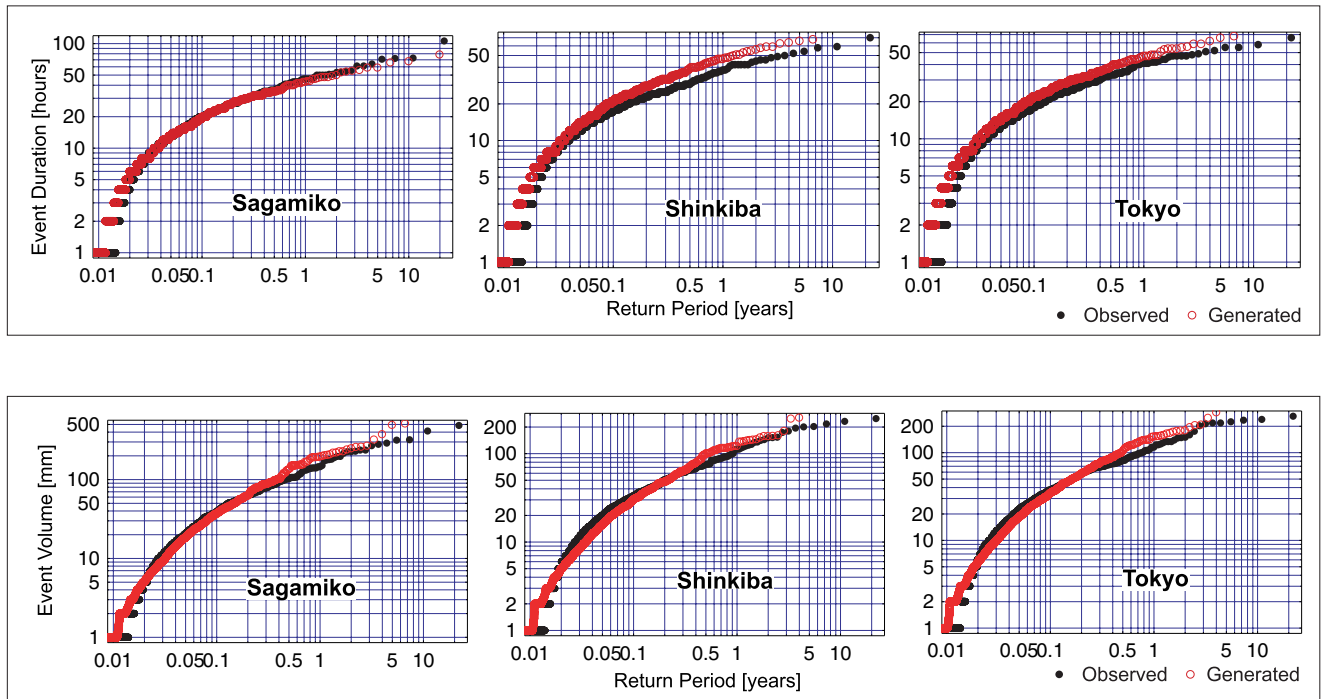


Fig. 13. Comparison of statistics of individual rain events. The convention that a continuous dry period of more than 6 h separates two rainfall events, was adopted to separate rainfall events from time series. Top: The durations of rain events. Bottom: The rainfall amount of rain events.

important. Rainfall events were identified using the rule that a continuous dry-stretch of more than 6 h separates one event from another. (There is no special reason to select the particular value of 6 h. However, the large drop in autocorrelations at initial lags and the prime importance of the rainfall structure at several hours in urban flood analysis, supports a value around this.) Both the total rainfall amount and the duration of the events had similar distributions in synthetic and observed data (Fig. 13).

To examine the model performance further in predicting extreme rainfall events, an analysis on an annual basis was performed. The largest annual hourly rainfall values were extracted from observed and synthetic hourly data and an extreme value distribution was fitted. Type III distribution was given by:

$$P(X \leq x) = e^{-[(x-\varepsilon)/(\theta-\varepsilon)]^k} \quad (10)$$

where  $\varepsilon$ ,  $\theta$  and  $k$  are constants. Figure 14 shows the plots for Tokyo and Funabashi stations on *log-extremal (Weibull)* probability paper (Haan, 1977). Estimated magnitudes of 10, 25, 50 and 100 year events are also given. Taking into consideration the widely accepted problems related to extrapolation of extremal fittings (Chow, 1964), it can be concluded that the proposed method reproduces the extreme events reasonably well.

## Conclusions

The multifractal properties of rainfall fields provide a means of relating rainfall distributions at various temporal scales. The widely adopted method of analysing multifractal properties of various fields does not work well for estimating higher resolution rainfall distributions from those observed at a daily resolution. This may be because of the unavailability of a sufficient scaling regime above the daily scale — a key requirement of all methods based on linear regression in log-log space. However, with a slight modification of the method of model fitting, it was possible to devise a method that performed better than the linear regression-based approaches.

To convert the ‘approximately equal’ relation (in Eqn. 1) to an exact equality, some functional form has to be introduced as a multiplier. In the present analysis, its simplest form, namely a constant, was assumed to be sufficient for that purpose. The results indicate that this is indeed a fair assumption for the purpose of predicting rainfall distributions at higher temporal resolutions than those at which they had originally been measured.

The disadvantage of this method of calculating the codimension function directly is that a value for a normalising constant,  $b$ , has to be calculated. However, as illustrated in Fig. 15, the sensitivity of the quality of estimation to small variations ( $\pm 10\%$ , for example) in the

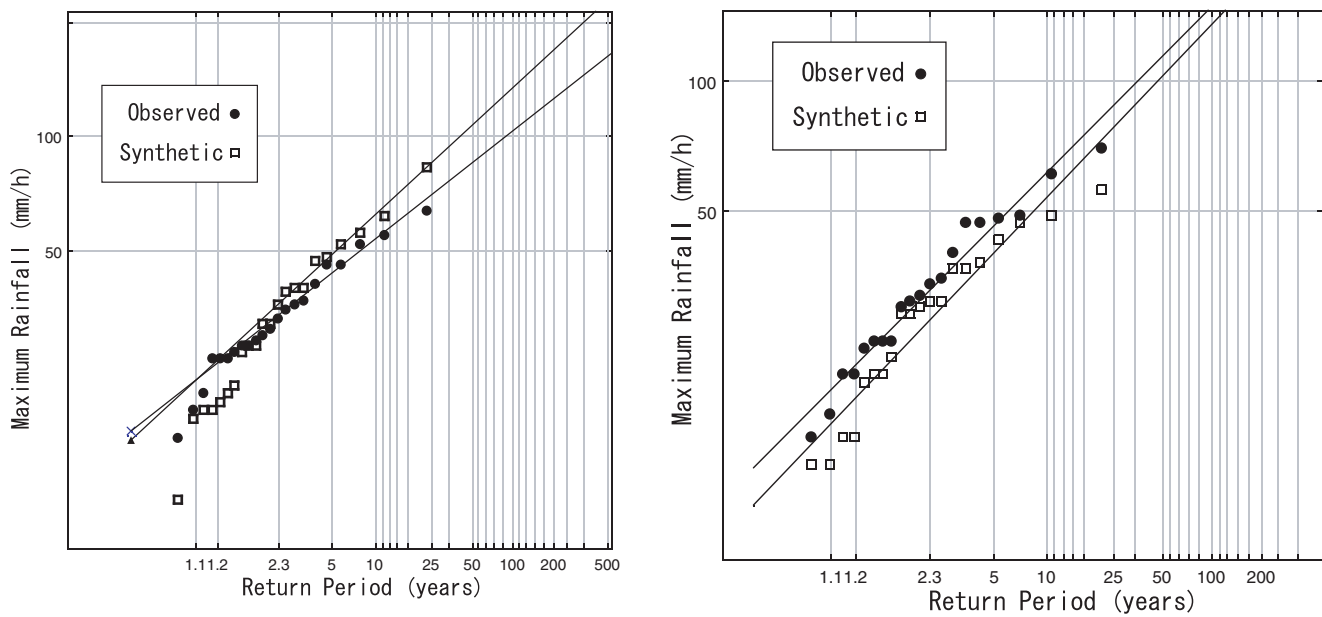


Fig. 14. Extremal distributions for Tokyo (Left) and Funabashi (Right) stations.

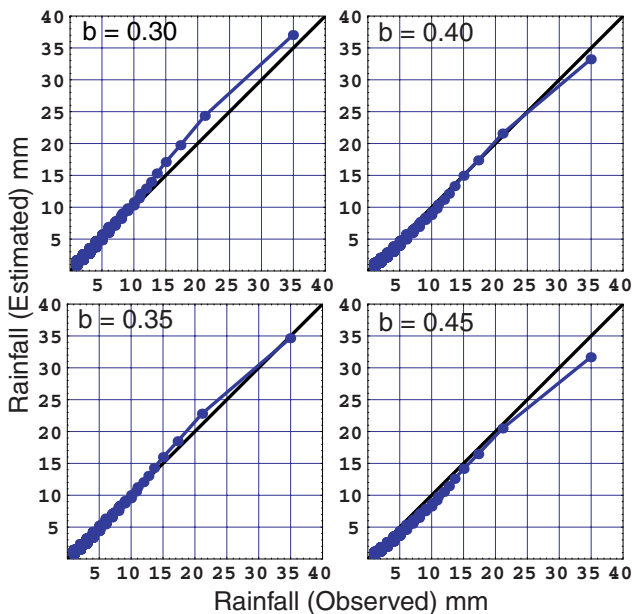


Fig. 15. The sensitivity of the high resolution rainfall estimation to small variations in  $b$ . (Funabashi).

value of  $b$  is not high. Thus, estimating an approximate value for the constant may be adequate for accurate prediction of high-resolution data if the same value is used to derive multifractal parameters and to generate rainfall values. However, using the trivial value of unity can neither be justified nor does it result in a reasonable estimation.

Previous research has established that the multifractal

parameters  $C_1$  and  $\alpha$  are not completely independent of each other. For example Pathirana (2001) found a negative correlation of about 0.41 between the two parameters in an analysis (using a double trace moment method proposed by Lavallée, 1991) involving about 1000 datasets from Japan. Further, when the available range of  $\gamma$  values is limited, (1) small changes in  $\alpha$  do not affect the shape of the codimension function and (2) very similar codimension curves can result from quite different pairs of  $C_1$  and  $\alpha$  values. These observations are equally true for the present analysis and, therefore, the individual values of  $C_1$  and  $\alpha$  should be used with care in comparison with different studies.

The general pattern observed in the quantile-quantile plots is that the estimation deviates gradually from that observed at high intensity rainfall values. Even so, the deviations would not generally exceed five to ten percent of the observed values for the last quantile (most extreme values) estimated. In a similar study of Canadian rainfall, Nguyen and Panday (1994) reported an opposite tendency. At present it is difficult to avoid this tendency consistently. Perhaps this disagreement at the extreme end may be the relatively low number of statistics for fitting the model at those intensities. Assuming that the codimension model can capture the variability at all intensities perfectly, one may argue that fitting the model using moderate and small intensities should get the extreme values correct. However, it is difficult at present to resolve the argument.

In addition to the intensities, several properties that are crucial to the use of rainfall series for hydrological studies

were used to compare the quality of the synthetic hourly data against the observed data. These included: temporal structure of rainstorm distribution, autocorrelations, statistics at the level of rainfall events and the return period of different intensities based on extreme value distributions. These tests indicated that the synthetic rainfall series mimic closely the observed hourly data, from a statistical point of view.

It can be concluded that it is possible to apply multifractal theories to derive hourly rainfall distributions from those observed at daily scale, using the direct calculation of the codimension function. It has been shown that it is possible to estimate hourly rainfall distributions using a daily dataset of about 20 years with only 1mm precision: such requirements which can be satisfied by operational rainfall observations available in many parts of the world.

## References

- Chow, V.T., 1964. *Handbook of Applied Hydrology*. McGraw-Hill, New York, USA.
- de Lima, M.I.P. and Grasman, J., 1999. Multifractal analysis of 15-min and daily rainfall from a semi-arid region in Portugal. *J. Hydrol.*, **220**, 1–11.
- Deidda, R., Benzi, R. and Siccardi, F., 1999. Multifractal modeling of anomalous scaling laws in rainfall. *Water Resour. Res.*, **35**, 1853–1867.
- Feder, J., 1989. *Fractals*. Plenum Publishing, New York.
- Frisch, U.P., Sulem, P.L. and Nelkin, M., 1978. A simple dynamical model of intermittency in fully developed turbulence. *J. Fluid Mech.*, **87**, 719–724.
- Grigoriu, M., 1995. *Applied Non-Gaussian Processes*. PTR Prentice Hall, Englewood Cliffs, NJ, USA. 442pp.
- Gupta, V. K. and Waymire, E., 1990. Multiscaling properties of spatial rainfall and river flow distributions. *J. Geophys. Res.*, **95**, 1999–2009.
- Haan, C.T., 1977. *Statistical Methods in Hydrology*. Iowa State University Press, Iowa, USA.
- Lavallée, D., 1991. *Multifractal analysis and simulation technique and turbulent fields*. PhD thesis, McGill University, Montréal, Canada. 133pp.
- Lavallée, D., Lovejoy, S. and Schertzer, D., 1991. On the determination of the codimension function. In: *Non-linear variability in geophysics: Scaling and Fractals*, S. Lovejoy and D. Schertzer (Eds.), Kluwer, Dordrecht, The Netherlands. 175–183.
- Lovejoy, S. and Mandelbrot, B.B., 1985. Fractal properties of rain, and a fractal model. *Tellus*, **37A**, 209–232.
- Lovejoy, S. and Schertzer, D., 1990. Fractals, raindrops and resolution dependence of rain measurements. *J. Appl. Meteorol.*, **29**, 1167–1170.
- Mandelbrot, B.B., 1974. Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier. *J. Fluid Mech.*, **62**, 331–350.
- Mandelbrot, B.B., 1967. How long is the coast of Britain? statistical self-similarity and fractional dimension. *Science*, **156**, 636–638.
- Meyer, R.R. and Roth, P., 1972. Modified damped least squares: An algorithm for nonlinear estimation. *J. Inst. Math. Appl.*, **9**, 218–233.
- Monin, A.S. and Yaglom, A.M., 1975. *Statistical Fluid Mechanics*. Vol. 2, MIT Press, Mass., USA. 874.
- Nguyen, V.-T. and Panday, G.R., 1994. Estimation of short-duration rainfall distribution using data measured at longer time scales. *Water Sci. Technol.*, **29**, 39–45.
- Novikov, E.A. and Stewart, R., 1964. Intermittency of turbulence and spectrum of fluctuations in energy-dissipation. *Nauk. SSSR. Geofiz.*, **3**, 408–412.
- Olsson, J., 1996. Validity and applicability of a scale-independent multifractal relationship for rainfall. *Atmos. Res.*, **42**, 53–65.
- Olsson, J., 1998. Evaluation of a scaling cascade model for temporal rainfall-disaggregation. *Hydrol Earth Syst. Sci.*, **2**, 19–30.
- Olsson, J., Niemczynowicz, J. and Berndtsson, R., 1993. Fractal analysis of high-resolution rainfall time series. *J. Geophys. Res.*, **98**, 23265–23274.
- Pathirana, A., 2001. *Fractal Modeling of Rainfall: Downscaling in Time and Space for Hydrological Applications*, PhD thesis, University of Tokyo, Tokyo, Japan. 148pp.
- Rosenbrock, H.H., 1960. An automatic method of finding the greatest or least value of a function. *Comput. J.*, **3**, 175–184.
- Schertzer, D. and Lovejoy, S., 1987. Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *J. Geophys. Res.*, **92**, 9693–9714.
- Svensson, C., Olsson, J. and Berndtsson, R., 1996. Multifractal properties of daily rainfall in two different climates. *Water Resour. Res.*, **32**, 2463–2472.
- Tessier, Y., Lovejoy, S. and Schertzer, D., 1993. Universal multifractals: theory and observations for rain and clouds. *J. Appl. Meteorol.*, **2**, 223–250.
- Wilson, J., Schertzer, D. and Lovejoy, S., 1991. Continuous multiplicative cascade model of rain and clouds. In: *Non-linear variability in geophysics: Scaling and Fractals*, D. Schertzer and S. Lovejoy (Eds.), Kluwer, Dordrecht, The Netherlands. 185–207.